Analyzing Students' Preferences for LLM-Generated Analogies

Seth Bernstein
Temple University
Philadelphia, PA, United States
seth.bernstein@temple.edu

Paul Denny
University of Auckland
Auckland, New Zealand
paul@cs.auckland.ac.nz

Juho Leinonen
Aalto University
Espoo, Finland
juho.2.leinonen@aalto.fi

Matt Littlefield
Temple University
Philadelphia, PA, US
matt.littlefield@temple.edu

Arto Hellas
Aalto University
Espoo, Finland
arto.hellas@aalto.fi

Stephen MacNeil
Temple University
Philadelphia, PA, US
stephen.macneil@temple.edu

ABSTRACT

Introducing students to new concepts in computer science can often be challenging, as these concepts may differ significantly from their existing knowledge and conceptual understanding. To address this, we employed analogies to help students connect new concepts to familiar ideas. Specifically, we generated analogies using large language models (LLMs), namely ChatGPT, and used them to help students make the necessary connections. In this poster, we present the results of our survey, in which students were provided with two analogies relating to different computing concepts, and were asked to describe the extent to which they were accurate, interesting, and useful. This data was used to determine how effective LLM-generated analogies can be for teaching computer science concepts, as well as how responsive students are to this approach.

KEYWORDS

large language models, computer science education, analogies

ACM Reference Format:
Seth Bernstein, Paul Denny, Juho Leinonen, Matt Littlefield, Arto Hellas, and Stephen MacNeil. 2024. Analyzing Students' Preferences for LLM-Generated Analogies. In *Proceedings of the 2024 Innovation and Technology in Computer Science Education V. 2 (ITiCSE 2024), July 8–10, 2024, Milan, Italy.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3649405.3659504

1 METHODOLOGY AND RESULTS

Previous research has indicated that analogies can facilitate the learning process by linking unfamiliar computer science concepts to real-world contexts, yet creating effective analogies can be a challenging and time-consuming process for both students and educators. Our study, conducted in an introductory C programming course at a large public university, involved 841 of 889 first-year Engineering students (95%). The course, emphasizing recursion, was part of the core curriculum. We created and tested two ChatGPT-generated analogies on recursion. Using a 5-point Likert scale, students evaluated their accuracy, usefulness, and interest. We analyzed the data using cumulative link mixed models (CLMM) and adjusted post hoc comparisons with the Tukey method to assess the educational impact of the analogies.

Analogy 1:
This program works like reading a book from the end to the beginning. You have to get to the last page (end of the string) first, and then you can start reading from back to front (each character printed after the recursive call).

Analogy 2:
Imagine a high school game where students stand in a line, and the goal is to say their name in reverse order, starting from the last student to the first. The teacher, Mr. C (representing our main function), asks the first student in line (the first character of our string) to start. Instead of immediately saying their name, the first student asks the student behind them to go first. This continues until the last student (the null character, '\0') is reached. This last student doesn't have anyone behind him/her, so instead of asking someone else, they just shout out their name. After the last student shouts their name, the one in front of them shouts, and so on, until we reach the first student again. In this code, each student represents a character of the word, and them waiting for the person behind to say their name represents the recursive call (reverseString(word + 1)). The action of them shouting out their name represents the printf function that prints the character.

Figure 1: ChatGPT-Generated Analogies for Evaluation

Our results show that analogy design significantly impacts perceived accuracy, engagement, and usefulness. Despite no errors in either, Analogy 2 was perceived as more accurate than Analogy 1 (Estimate = -1.21, $p < .01$), likely students perceive less detailed descriptions as less accurate. Additionally, Analogy 2 was found to be more engaging (Estimate = -1.90, $p < .01$) and useful (Estimate = -1.64, $p < .01$) compared to Analogy 1, highlighting its enhanced interest and benefits to students. Our exploration showed that varying the detail level of prompts influenced student ratings, with more detailed analogies being preferred by students.

Analogy 1	Accurate (%)	Interesting (%)	Useful (%)
Strongly Agree	8.5	4.1	6.9
Agree	49.7	23.4	27.6
Neutral	24.9	33.6	30.5
Disagree	14.9	29.3	26.1
Strongly Disagree	2.1	9.6	8.9
Analogy 2	**Accurate (%)**	**Interesting (%)**	**Useful (%)**
Strongly Agree	29	23	29.5
Agree	49.4	48.7	42.5
Neutral	15.1	19.7	19.1
Disagree	6.3	6.7	6.8
Strongly Disagree	0.2	1.9	2.2

Table 1: Responses to Analogies Assessment

ITiCSE 2024, July 8–10, 2024, Milan, Italy

ACM ISBN 979-8-4007-0603-5/24/07.
https://doi.org/10.1145/3649405.3659504